DOCUMENT RESUME

ED 401 591 CS 509 385

AUTHOR Fenton, Ray; O'Leary, Neil
TITLE Assessment Issues: Speech Communication Association. Speaking. Listening, and Media Literacy Standards.
PUB DATE Nov 96
NOTE 18p.; Paper presented at the Annual Meeting of the Speech Communication Association (82nd, San Diego, CA, November 23-26, 1996). Prepared for Speech Communication Association Committee on Assessment and Testing.
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Standards; *Audiolingual Skills; *Communication Skills; Curriculum Development; Elementary Secondary Education; *Evaluation Methods; Higher Education; *Student Evaluation; *Testing; Test Reliability; Test Validity
IDENTIFIERS Communication Behavior; Media Literacy; *Speech Communication Association

ABSTRACT

As individual states and communities work out curriculum and instructional expectations, high quality assessments of communication skills will continue to evolve. An active cycle of setting goals, developing instruction, teaching, testing, and assessing student and system success is the key to the process of renewal and improvement. All types of information needed to guide decisions build upon and must be grounded in successful student demonstrations of what students know and are able to do. Student communication behaviors must be approached from a number of perspectives using multiple indicators to make reasonable, informed judgments. In the United States, standards and expectations for student performance are currently being established at the state level. The actual responsibility for assessing student performance falls to the local education agency. Additional help is available from the national Speech Communication Association (SCA) which has published national standards in speaking, listening, and media literacy for four age ranges: K-3, 4-6, 7-8, and 9-12. The need to gain insight into the relative performance of individuals from time to time and group to group adds complexity to the assessment process. The primary source of technical complexity for the assessment of human communication is that measures of performance require some form of human judgment. The single general scale used by the National Assessment of Educational Progress or the multiple scales tied to specific tasks such as those used in Alberta, Canada, and Kansas serve as examples of qualitative assessments of student performance. Use of published instruments can save money and often have proven validity and reliability, but it is unlikely that published instruments will exactly fit locally developed standards for student performance. (Contains 20 references and 4 tables.) (RS)

ED 401 591

Assessment Issues:

Speech Communication Association

Speaking, Listening, and Media Literacy

Standards

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

R. Fenton

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Ray Fenton
Neil O'Leary

Prepared for Speech Communication Association
Committee on Assessment and Testing
Speech Communication Association
Annual Convention
San Diego, California
November, 1996

CS505385

BEST COPY AVAILABLE

As in most important human enterprises, maintaining a high quality communication education program requires a vision of the goal to be reached, commitment to goal attainment, and the application of effort. An active cycle of setting goals, developing instruction, teaching, testing, and assessment of student and system success is the key to the process of renewal and improvement. Each element in the cycle is itself a process that must be given attention, time, and support.

Developing, administering, and reporting assessments is a key element in the process of renewal and improvement. Assessment provides the feedback that is needed to judge goal attainment and to focus efforts to improve performance.

Most states now have language arts standards which call for the development of communication skills. Many provide support through the provision of model curriculum frameworks and a few support assessment of communication skills.

Local State Education Agency and state communication association are the first places to contact when a school or school district sets out to improve communication programs. Additional help is now available from the national Speech Communication Association (SCA) which has published national standards in speaking, listening, and media literacy as well as a model curriculum framework and the recommendation that there be performance benchmarks set for communication competency at four age ranges: K-3, 4-6, 7-8, and 9-12.

While it is not likely that any system of benchmarks and competencies has been defined which meets all of the needs for assessment at the various levels of the K-12 educational system, there has been a great deal of good thinking and many high quality assessment instruments have been developed. Published instruments and the suggestions provided by State Education Agencies may serve as guides.

This paper briefly examines factors which must be considered in the development of systems for the assessment of communication skills and some of the sources that are already available to developers of assessments and assessment systems. It is meant to be a starting place and aid to those who desire to assess and monitor student performance as part of a system of instructional renewal and improvement.

Assessment and Decision-Making

The chart below provides an overview of the various types of information needed for effective renewal and improvement at various levels of the educational planning and decision-making process. Each level has its own responsibility within our educational system. While it sometimes seems that the type of information needed to guide decisions at one level has little utility for the other levels of the system, it is actually the case that all levels build upon and must be grounded in successful student demonstrations of what students know and are able to do. The overriding characteristic of effective and successful systems is that they allow and

generate the credible information needed by all levels of decision makers without undue cost or duplication of activities.

Table I

Systemic Assessment Needs
Good Measurement/Good Measures
A Sense of Purpose

Level of Review Level of Data	Decision Makers Focus of Decisions	Types of Decisions Level of Impact
Policy Level and Program Accountability	Public President/Secretary of Education Congress/State Legislature Governor/Chief State School Officer Superintendent/School Board	National Policy State and Local Goals Expected Systemic Outcomes Resource Allocation
Management and Program Improvement	State Education Department School District Administration Content Area Specialists	Content Standards Instructional Frameworks Large Scale Assessments Resource Allocation
Goal Setting Instructional Planning Performance Reviews	School Level Administrators School Governance Groups Teachers	Local Goal Attainment Resource Allocation School and Grade Level Planning Class Group Instructional Decisions
Personal Growth and Development Goal Setting Need Assessment	Students	Grades and Assessments Diagnosis Individualized Instructional Decisions Motivation Self-Direction

An authentic, high quality system of assessment collects valid information consistent with the classroom environment and useful for students and teachers (Wiggins, 1993). Doing authentic assessment and providing good information for making decisions at all of the levels of the hierarchy of educational planning and accountability is the challenge that must be accepted by all of those who make a commitment to systematic, systemic improvement.

Areas for Assessment

Student success is educational system success. Competent communication is complex and student communication behaviors must be approached from a number of perspectives. This demands the use of multiple indicators to make reasonable, informed judgments. Use of multiple indicators help to assure that the communication skills sampled are a sufficient representation of individual behavior for fair judgments of individuals and adequate for making instructional decisions.

Competent performance as a communicator is the product of a desire to communicate (affect and attitude), a realization of what can be done under a given set of circumstances (cognition and knowledge), and the ability to act effectively (skill and behavior). Some types of assessment are more effective than others in answering key questions about student success. The table below indexes the various types of common communication assessments and the primary questions that they answer. Defining the specific questions which an assessment system must answer is the first step in building an effective assessment system.

Table II
Elements of Competent Communication
Primary Assessment Types and Defining Questions

Area	Types	Source	Defining Questions
Affect and Attitude	Self-Report	Teacher Constructed	How do I feel?
	Oral Interview	Teacher Constructed or Scripted	How does the student feel?
	Performance Rating Scale	Standardized	How do I my feelings and predispositions compare?
Cognition and Knowledge	Self-Report	Teacher Constructed	How did I do relative to instruction?
	Oral Interview	Teacher Constructed or Scripted	What does the student know?
	Performance Rating Scale	Standardized •Criterion Referenced •Norm Referenced	How does my knowledge compare? •Does knowledge meet the expected standard? •Is knowledge comparable to knowledge of members of a reference group?
Skills and Behaviors	Self-Report	Teacher Constructed	Do I fulfill the expectations set for of assignment?
	Oral Interview	Teacher Constructed or Scripted	How well does the student perform?
	Performance Rating Scale	Standardized •Criterion Referenced •Norm Referenced	How well do I fulfill the standard? •Do skills and behaviors meet the expected standard? •Are skills and behaviors comparable to those of members of a reference group?

According to the National Education Standards and Improvement Council. it is common in other countries for the government to establish nationwide educational standards and the important questions to be answered. In China, for example, the standards are set by the State Education Commission in Beijing. In Japan, it is the Ministry of Education in Tokyo (Manibushi). In England, it has been a national responsibility since the Educational Reform Act of 1988. These countries support and to some extent set performance standards through their national systems of examinations, syllabi, and guides for scoring examinations [National Education Standards and Improvement Council (1993)]. In the United States, standards and expectations for student performance are currently being established at the state level. The actual responsibility for doing assessment of student performance is sometimes retained at the state level but is often the responsibility of the local education agency.

The effort required to directly observe and assess student performance that is inherent in communication education has limited the entry of state education agencies into the direct assessment of student communication competence. Direct assessment of speaking, listening, and media competencies has generally been left to local education agencies. This is often within the context of specific classes or curriculum areas but the form of assessment is often not specified. While some states define the specific behaviors to be assessed or provide assessments, it is more often the case that the state specifies general goals, benchmarks, and reporting procedures.

The transfer of the responsibility for actual assessment to local education agencies, schools and classrooms provides an opportunity to tailor assessment to instruction. This is important because there is always a tension created when individual student assessments have to serve more than one purpose and answer questions that relate to multiple levels of decision making.

When assessments are defined to serve the purposes of the classroom, it is more likely that there will be a collection of information which ties directly to the teacher-student relationship. In some cases, actual performances may even be recorded and become part of a student record in the form of a sound recording, video, or computerized portfolio. As one moves up the ladder of generalization and away from the classroom, data is often reduced into a form which characterizes some attribute of the attitude, knowledge, or performance. When data is reduced into a form which can be used to characterize the performance of groups and provide insight into the relative levels of accomplishment over time or among groups, technical considerations become critical.

Table III
Uses of Assessment Information

Primary Level of Use	Type of Information	Level of Review
National and State Policy Setting Standards Assessment	Comparative Summaries Norm Referenced Criterion Referenced	Broad Goal or Standard assessed relative to performance criteria or prior performance
Community and School District Management Program Assessment	Comparative Summaries Norm Referenced Criterion Referenced	Broad Goal or Standard assessed relative to performance criteria and prior performance for grade levels programs, and curriculum
Classroom Goal Setting Instructional Planning Performance Review	Comparative Summaries Norm Referenced Curriculum Referenced Skills Referenced	Compares performance of specific classroom groups of students with what is expected based on locally set criteria
Individual Student Assessment Personal Achievement Individual Growth	Curriculum Referenced Skills Referenced Mastery	Compares performance of the individual to what is expected based on standards and exposure to instruction

Technical Concerns

The need to gain insight into the relative performance of individuals from time to time and group to group adds complexity to the assessment process. Fair comparisons demands reliable and valid measurement. No assessment that is not valid or reliable can be fair to individuals or to programs.

The technical aspects of assuring fairness have been discussed in a number of sources. The most directly related to assessment of communication skills is the Speech Communication Association's own guide to standardized assessment instruments, Large Scale Assessment of Oral Communication: K-12 and College (Morreale and Backlund, 1996). Other standard references extend the discussion to large scale testing programs. Some of these include the CRESST guide, Improving Large-Scale Assessment extends the discussion to consider the choices which must be made when developing or adopting a large scale system (Aschbacher and Baker, 1988). And, the American Educational Research Association and American Psychological Association's standards which assure fair and equitable treatment for test takers (Committee to Develop Standards for Educational and Psychological Testing, 1985). If all the requirements set out in these guides are met, it is likely that the assessments will be of acceptable quality and have the credibility needed to influence decision making.

The primary source of technical complexity in system for the assessment of human communication is that measures of performance require some form of human judgment. To be fair for individuals, assessment systems must be standardized and objective to yield the same demands on those assessed and the same score for similar performances.

In the language of testing, "standardized" refers to strict adherence to the same procedures in test administrating and scoring. Standardization produces similar tasks and provides an equal opportunity for performance. The goal of standardization is to minimize the differences in performances and scores which might arise from differences in procedures from place to place and time to time.

"Objective" is also applied both to tests and scoring. Objective tests are closely tied to the behaviors which are critical to demonstrations of competence and knowledge. The design of the test and the task assigned are critical in the production of objective tests. In the process of assigning qualitative ratings or judging the correctness of answers, "objective" refers to the direct link between criteria for judgment and the resulting score. Objective scoring goes beyond the mere consistency of scoring (reliability) that is required in any fair assessment so that each score links in a logical way to a desired outcome. Scores on a good objective test are meaningful representations of competence.

Another important consideration which may go beyond the common conception of technical factors is cost. Good assessment that provides meaningful information is expensive. It is almost always thought of as too expensive if it is treated as an add-on to instruction rather than a critical element in the process of instruction and improvement.

Lawrence Pincus of the UCLA Center for the Study of Evaluation set out a conceptual framework for analyzing the costs of performance assessment (1994). The direct costs are development, production, training for administration, training for interpretation of results, training in related instruction, test administration, management, scoring, reporting, and program evaluation. The greatest cost to be considered, however, is not the monetary cost of the testing itself. It is the cost to students and teachers in time taken away from instruction and regular classroom activities. These costs make it incumbent on those who develop assessment systems to take care when considering the functions of the assessment system and the credibility of the system in providing answers to questions about the adequacy and efficacy of the instruction in preparing competent communicators.

The technical consideration is how much behavior must be assessed to make a fair judgment.

The costs of assessment often require that assessments be narrowly focused and only measure the most important skills. Time and money sometimes seem to dictate assessments that lack the spontaneity and flexibility of day–to–day classroom activity. In the classroom, teachers can assess and respond immediately with critical judgments and instructional interventions. The well–made standardized and objective measure must retain some of these qualities, provide a positive climate for learning, and be of obvious value to students and teachers to be accepted. The challenge is to retain the opportunities for valuable, critical judgments and instructional interventions by teachers without sacrificing the required objectivity and fairness required for accountability.

Table IV
Scoring Types

Primary Level of Use	Type of Information	Level of Review
National and State Policy Setting Standards Assessment	Comparative Summaries Norm Referenced Criterion Referenced	Standardized Objective
Community and School District Management Program Assessment	Comparative Summaries Norm Referenced Criterion Referenced	Standardized Objective
Goal Setting Instructional Planning Performance Review	Comparative Summaries Norm Referenced Curriculum Referenced Skills Referenced	Standardized Objective
Student Assessment	Comparative Individual Curriculum Referenced Skills Referenced Master	Standardized Objective Individualized Critical

The lack of strong correlation between grades given by teachers, short answer measures of affect and knowledge, and performance assessments makes it difficult to use only one type of measure to represent the mastery a student has of any academic area which emphasizes both the development of positive predispositions, knowledge, and performance skills. While the face validity and authenticity of performance assessments make them the favored choice as instruments to guide instruction and impact student performance (Wiggins, 1993a, 1993b), there is often a difficulty in representing what a student knows and what a student can do with only one type of assessment (Popham, 1983, 1994).

There are also notable differences in student responses to seemingly similar problems. The low correlations between performances on seemingly similar assignments may be due to contextual differences as suggested by Wiggins (1993a). Or there may be even deeper reasons related to the cognitive complexity demanded by various tasks. Explorations in the domain of science where factual knowledge, procedural skill, and the ability to apply knowledge and skill in unique situations

parallel the requirements of communication have prompted a number of investigations of this problem. The landmark study by Shavelson, Baxter, and Gao (1993) suggests that there are a number of different factors which might affect the ability to respond to different tasks:

- tasks might not elicit the same cognitive skill,
- tasks may call upon a different level of cognitive knowledge, or
- data collection and the scoring procedures may be more sensitive in some cases to task-specific knowledge (content and process knowledge).

Whatever the cause of the commonly found lack of correspondence between measures of knowing and doing and inconsistency in performances on seemingly similar tasks, the lack of correspondence is of particular concern in the area of communication where both competent performance and increased knowledge are the hallmarks of student success.

Reporting Performance

One of the most powerful outcomes of the assessment system is the ability to rate performance along a unified scale of communication competence which is grounded in the important elements of attitude, knowledge, and performance expected at each benchmark grade or age level the value of such scales can be seen in the scale developed by the National Assessment of Educational Progress to characterize student performance in various academic areas. While the NAEP efforts to apply sound psychometric methodology to ground the scales on fair assessments from multiple performances are complex, the scales themselves give clear and understandable descriptions of student accomplishment (Campbell et al., 1994).

National Assessment of Educational Progress
Proficiency Scale

Below Basic –

Basic – This level denotes a partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.

Proficient – This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real world situations and analytical skills appropriate to the subject matter.

Advanced – This level signifies superior performance.

Scale points may also be connected to critical descriptions of behavior which characterize specific communication performances. A model guide developed by the Kansas Speech Communication Association Task Force on Speaking and Listening assessment does this through setting criteria for Speaking and Listening Assessment. It provides model assessments and scales for listening, group participation, and speaker-to-audience performance for elementary, middle school, and high school age students. Critical observable elements are identified for each type of performance and set out in descriptive rubrics. For example, the four critical features of participation in small group discussion for elementary students are identified as "Listens Actively to Others", "Participates Effectively in Discussion", "Demonstrates awareness and Sensitivity to Conversational Partners", and "Demonstrates Awareness and Concern for Accomplishing Group Goals and Tasks". The ability of the student to "Participate Effectively in Discussion" is judged through the observation of performance and scoring on a five-point scale which is grounded in a rubric based on descriptions of performance. Scores range from one to five. Three points are described:

Kansas State Speech Association Discussion Scale

> 1. Contributions are minimal or nonexistent. When contributes uses language that is vague, inaccurate or offensive. Does not speak clearly and/or loudly enough to be understood. Does not have information to share. It is difficult to follow their contributions. Their comments are not relevant to the discussion taking place.
>
> 3. Contributions promote discussion. Uses language that does not offend and helps to create mutual understandings of the issues under discussion. Speaks clearly and loudly enough to be understood. Has information to share that is beyond personal opinion such as examples from other students, parents, teachers, magazines, books, or t.v. shows. Their comments are easy to follow, make sense, and are relevant to the discussion taking place.
>
> 5. Contributions are valuable to the outcome of the group's discussion. Uses language which enhance the mutual understanding of the issues under discussion. Refers to information that is specific and concrete: such as illustrations, examples, or comparisons. Comments are exceptionally clear and stimulate the discussion. (Kansas Speech Communication Association, 1994, p. 9)

Each of the other elements is scored on a similar scale, making for a total of four scores on four elements of group discussion.

A similar approach is included in the Oral Communication Evaluation: English 30/33 manual prepared by Alberta Education, which provided ratings on a 1 to 5 scale for Interaction, Comprehension Strategies, and Performance Skills to be demonstrated in a group discussion assignment (Alberta, 1990).

Alberta Education Discussion Scale

5. Proficient: the student handles the task consistently, frequently exceeding expectations. Thoughtful and substantial ideas and details are presented. Clarifying and elaborating occurs as meaning is extended for self and others.

4. Capable: The student handles the task thoroughly, fulfilling expectations. Clear ideas and substantial details are presented. Probing for meaning occurs.

3. Adequate: The student handles the demands of the task mechanically, fulfilling the surface requirements. Appropriate ideas and adequate details are presented. The student may be satisfied with quick answers, lacking the curiosity or security to probe for meaning of more substance.

2. Limited: the student needs continual clarification of expectations of the task. :Limited ideas and details are presented. The student rarely probes for meaning.

1. Poor: The student is consistently unable to handle the task. Confusion about the task results in the presentation of scant ideas and details. The student almost never probes or comments on contributions, so there is little demonstration of the pursuit of understanding.

The single general NAEP scale or the multiple scales tied to specific tasks such as those used in Alberta and Kansas for assessing discussion all serve as examples of qualitative assessments of student performance which may generate ratings useful in examining the relative efficacy of programs. The grounding of the Kansas and Alberta scales in specific descriptions of the behaviors to be demonstrated by competent group members place them much closer to the classroom than more general descriptive scales of NAEP.

More specific scales make better guides for teachers and curriculum developers. But this occurs at the expense of having a simple single number which may be useful to a State or School District which wishes to present a unidimensional scale of program quality. More general scales may make it easier to construct programs and curriculum elements. but this occurs at the expense of direct links to instructional practice. Careful consideration has to be given to the scales used to measure performance.

Selection of Instruments

Finally, there is a paradox in the selection of instruments. Use of published instruments can save money on instrument development. Published instruments

often have proven validity and reliability for certain known uses which gives them credibility. However, it is not likely that published instruments will exactly fit locally developed standards or benchmarks for student performance. The chief advantage of using existing instruments is that they have some history of use which may demonstrate that they may be successfully utilized to assess similar goals. A second advantage is that some have been used with enough students that normative information is available. Unfortunately, there are few performance measures which offer normative information and have demonstrated validity and reliability with k-12 students (Morreale and Backlund, 1996).

A number of published assessments are described in the Speech Communication Association publication, Large Scale Assessment of Oral Communication: K-12 and College. Some examples are included below. Additional models may be found through the national Speech Communication Association, State communication associations, State Departments of Education, regional educational laboratories, the ERIC Center on Testing and Assessment, and the Center for Research on Evaluation and Testing at UCLA. The Speech Communication Association Committee on Assessment and Testing is addressing the issues of standards and assessment. Measures are being developed to examine student competencies in a variety of areas. The SCA Committee on Assessment and Testing may be reached through the Speech Communication Association National Office, (703) 750-0533.

Two short instruments which focus on the observation of specific student behaviors related to general communication skills. These may serve as models for the assessment of skills in young children which are useful in a school setting. The Joliet 3-minute Speech and Language Screen and the California Achievement Tests (CAT 5) Listening and Speaking Checklist are different in purpose and in assessment approach. The Joliet screen is a quick classification of students based on the ability of the student to respond to a standard conversational prompt in a dyadic interview (Communication Skill Builders, 1993). The CAT 5 Assessment asks the classroom teacher to watch students and listen to them during day-to-day classroom activities. The teacher prepares a report which may be shared with parents based on three observations of students in situations where they might exhibit behavior related to Listening, Listening Comprehension, Critical Listening, Speaking, and Group Participation. Students also consider and rate their own behavior. Teacher observations and student self-reflection are both based on a judgment of whether a specific behavior such as "Is considerate of others' ideas and feelings" takes place "Most of the time", "Sometimes", or "Not Usually" (CTB Macmillan/McGraw-Hill, 1993).

A model secondary instrument based on a number of the Speech Communication Association Speaking and Listening Competencies (1994) is the Communication Competency Assessment Instrument: High School Edition, which is a "rating instrument designed to assess high school students' communication readiness for upper-level course work: their abilities to speak and relate to others both within and outside classroom contexts . . ." (Ruben, 1993, p. 1). Each student gives a three-minute persuasive talk which may be video taped and is asked to respond to statements about high school experiences in a dyadic interview. Each student assessment takes about half an hour and results in fifteen ratings.

Areas Rated on the Communication Competency Assessment High School

1) Pronunciation,
2) Tone of Voice,
3) Speech Clarity,
4) Persuasiveness,
5) Clarity of Ideas,
6) Ability to Express and Defend Ideas,
7) Ability to Recognize Non-understanding,
8) Ability to Introduce Self,
9) Asking a Question,
10) Answering a Question,
11) Expressing Feelings,
12) Use of Chronological Order,
13) Ability to Give Directions,
14) Ability to Describe Another's View, and
15) The Ability to Describe Differences in Opinion About the Steps Necessary to Accomplish Academic or Vocational Goals.

Careful attention to consistent directions, clear prompts, and systematic training of raters makes the CCAI-H.S. and performance-based assessment which has demonstrated high reliability (Rubin, 1993).

There are not published instruments which address all of the Speech Communication Association Listening, Speaking and Media Literacy standards or the grade appropriate competencies which have been identified in the much more specific Speech Communication Associations K-12 Curriculum Guidelines. As David Cohen observed in his article on national standards in the Phi Delta Kappan, "Even the language of standards is still wet behind the ears, barely sketched out." (p. 755). This is doubly true in regard to assessment. We are only starting to understand the potential impact of standards on the improvement of curriculum and programs. As individual states and communities work out curriculum and instructional expectations, high quality assessments which may be used to assess student and program success will continue to evolve.

References

Alberta Education, Oral Communication Evaluation: English 30/33 Activities and Scoring Guides. Edmonton, Alberta: Education Alberta.

Aschbacher, Pamela and Baker, Eva (1988), Improving Large-Scale Assessment. Los Angeles, CA: UCLA Center for Research and Evaluation, Standards and Student Testing (CRESST).

CTB Macmillan/McGraw-Hill (1993) , Listening and Speaking Checklist. Monterey, CA: CTB Macmillan/McGraw-Hill.

Campbell, J.R., Donahue, P.L.O., Reese, C.M. and Phillips, G.W. (1996), NAEP 1994 Reading Report Card for the Nation and States: Findings from National Assessment of Educational Progress. U.S. Department of Education, Office of Educational Research and Improvement .

Cohen, David (1995), "What Standards for National Standards." Phi Delta Kappan, 76,10, Pp.

Committee to Develop Standards for Educational and Psychological Testing (1985), Standards for Educational and Psychological Testing, Washington, D.C.: American Psychological Association.

Jablon, J.R., Ashely, L.A., Marsden, D.B., Meisels, S.J., and Dichtelmiller, M.L. (1994), Omnibus Guidelines: Kindergarten through Fifth Grade, 3rd. Edition. Ann Arbor, MI: Rebus Planning Associates, Inc.

Kansas Speech Communication Association (1994), Speaking and Listening Assessment Guidelines.

Morreale, S.P., and Backlund, P.M. (1996), Large Scale Assessment of Oral Communication: K-12 and College, 2nd Ed.

National Education Standards and Improvement Council (1993). Promises to Keep: Creating High Standards for American Students. Report to the National Educational Goals Panel. Washington, DC: National Goals Panel.

Pincus, L. O. (1994). A Conceptual Framework for Analyzing the Costs of Alternative Assessment. Los Angeles, CA: CRESST/University of Southern California.

Popham, W.J. (1984). A guide to criterion-referenced test construction. Baltimore, MD: the Johns Hopkins University Press.

Popham, W.J. (1993). Educational Testing in America: What's right, what's wrong?: A criterion-referenced perspective. Educational Measurement: Issues and Practices, 12 (1) 11-14.

Rubin, R.B. (1993) Communication Competency Assessment Instrument: CCAI-High School Edition. Johnson City, TN: Spectra Inc. Publishers.

Shavelson, R.J., Baxter, G.P., and Gao, X. (1993). Sampling variability of performance assessments. Journal of Educational Measurement, 30(3), 215-232.

Speech Communication Association (1994). Speaking and Listening Competencies. Annandale, VA, Speech Communication Association.

Speech Communication Association (1996). Speaking, Listening, and Media Literacy Standards for K through 12 Education. Annandale, VA, Speech Communication Association.

Speech Communication Association (In Press). K-12 Curriculum Guidelines. Annandale, VA, Speech Communication Association (ISNM:0-944811-15-9).

Wiggins, Grant. (1993a). Assessment: Authenticity, context, and validity. Phi Delta Kappan, 75 (3), 200-214.

Wiggins, Grant (1993b). Assessing Student Performance: exploring the purpose and limits of testing. San Francisco, Jossey-Bass Publishers.